



Good Practice Guide No. 71

## The Measurement of Mass and Weight



ADVANCED  
TECHNICAL LEVEL

# The National Physical Laboratory (NPL)

NPL is the UK's National Measurement Institute, and is a world-leading centre of excellence in developing and applying the most accurate measurement standards, science and technology available.

NPL's mission is to provide the measurement capability that underpins the UK's prosperity and quality of life.

© NPL Management Limited, 2026

Version 1.0 – June 2004

Version 2.0 – April 2026

<https://doi.org/10.47120/npl.mgpg71>

## NPL Authors and Contributors

STUART DAVIDSON

PERDI POTGIETER

DAVIT GABELAIA

Find out more about NPL measurement training at [www.npl.co.uk/training](http://www.npl.co.uk/training) or our e-learning Training Programme at [training.npl.co.uk/](http://training.npl.co.uk/)

NPL made every effort to ensure all information contained in these Good Practice Guides was correct at time of publication. NPL is not responsible for any errors, omissions or obsolescence, and does not accept any liability arising from the use of these Good Practice Guides.

National Physical Laboratory  
Hampton Road  
Teddington  
Middlesex United Kingdom

**Telephone:** +44 (0)20 8977 3222  
**e-mail:** [training@npl.co.uk](mailto:training@npl.co.uk)  
[www.npl.co.uk](http://www.npl.co.uk)

*On the cover: Kilogram 18, the UK's national copy of the International Prototype Kilogram, in its storage container (top left); OIML shaped weight being put into a mass comparator (bottom left); NPL's next-generation Kibble balance, a primary realisation of the SI unit of mass (right).*

## **Abstract/Foreword**

This Good Practice Guide is intended as a useful reference for those involved in the practical measurement of mass and weight.

## **Acknowledgements**

The authors acknowledge the financial support provided by the National Measurement System Programme of the UK Department for Science, Innovation and Technology for Version 2 of this Good Practice Guide, and by the Department of Trade and Industry for Version 1.

# Contents

Introduction .....	1
1.1 The definition of the unit of mass .....	2
1.2 Realisation of the unit of mass .....	2
1.3 Kilogram consensus value .....	3
1.4 National mass standards .....	4
Weights .....	5
2.1 Material .....	6
2.2 Surface condition .....	6
2.3 Magnetism .....	7
2.4 Weight cleaning .....	7
2.5 Weight handling .....	7
2.5.1 Gloves .....	8
2.5.2 Tweezers .....	8
2.6 Clean surfaces .....	9
2.7 Weight storage .....	10
Weighing techniques .....	11
3.1 Direct reading measurements .....	12
3.2 Weighing by differences .....	12
3.3 Substitution weighing .....	13
3.4 ABA... calibration .....	14
3.5 ABBA calibration .....	14
3.6 Cyclic weighing .....	15
3.7 Weighing by sub-division .....	16
3.8 Calibration intervals .....	17
3.9 Scale errors and the use of makeweights .....	18
Air density measurement and buoyancy correction .....	19
4.1 True mass .....	20
4.2 Conventional mass .....	20
4.3 Buoyancy correction .....	20

4.4	Determination of air density from parametric measurements .....	22
4.5	The application of buoyancy corrections .....	22
Balance assessment .....		25
5.1	Balance location .....	26
5.2	Measuring equipment .....	26
5.3	Two-pan balances .....	27
5.4	Single-pan mechanical balance .....	28
5.5	Single-pan electronic balances.....	28
5.6	Assessment of two-pan balances .....	29
5.6.1	Routine periodic checks .....	29
5.6.2	Rider weight .....	29
5.6.3	Internal balance weights .....	29
5.6.4	Linearity of scale .....	29
5.6.5	Sensitivity .....	29
5.6.6	Repeatability of reading .....	30
5.6.7	Repeatability of measurement .....	30
5.7	Single-pan balances .....	30
5.7.1	Visual inspection and mechanical check .....	30
5.7.2	Drift.....	30
5.7.3	Calibration of internal weights.....	31
5.7.4	Effect of off-centre loading (eccentricity) .....	31
5.7.5	Hysteresis .....	31
5.7.6	Scale sensitivity (scale value) .....	31
5.7.7	Scale linearity .....	31
5.7.8	Repeatability of reading .....	31
5.7.9	Repeatability of measurement .....	32
5.8	Assessment of electronic balances .....	32
5.8.1	Hysteresis .....	33
5.8.2	Effect of off-centre loading (eccentricity) .....	33
5.8.3	Scale error and linearity .....	33
5.8.4	Repeatability of reading .....	33

5.8.5	Repeatability of measurement .....	34
5.9	Periodicity of calibration .....	34
Uncertainty in mass calibration .....		35
6.1	Introduction .....	36
6.2	The measurement .....	36
6.3	The uncertainty budget .....	36
6.3.1	Symbol .....	37
6.3.2	Sources of uncertainty .....	37
6.3.3	Values .....	38
6.3.4	Probability distributions .....	39
6.3.5	Divisor .....	41
6.3.6	Sensitivity coefficient ( <b><i>ci</i></b> ) .....	41
6.3.7	Standard uncertainty in units of measurand, <b><i>ui(Wi)</i></b> .....	42
6.3.8	Degrees of freedom, <b><i>vi</i></b> .....	42
6.3.9	Adding it all up .....	42
6.4	Air buoyancy uncertainty budget .....	44
6.5	Reporting the results .....	46
References .....		47



## Chapter 1

# Introduction

- The definition of the unit of mass
- Realisation of the unit of mass
- Kilogram consensus value
- National mass standards

## 1.1 The definition of the unit of mass

The SI unit of mass is the kilogram. In 2019 the definition of the kilogram was revised from a definition based on the value of a physical artefact, the international prototype of the kilogram (IPK), to a definition based on a fundamental constant of nature.

“The kilogram, symbol kg, is the SI unit of mass. It is defined by taking the fixed numerical value of the Planck constant  $h$  to be  $6.62607015 \times 10^{-34}$  when expressed in the unit J·s, which is equal to  $\text{kg}\cdot\text{m}^2\cdot\text{s}^{-1}$ , where the metre and the second are defined in terms of  $c$  and  $\Delta\nu_{\text{CS}}$ .”

The difference between ‘mass’ and ‘weight’ is that mass is a measure of the amount of material in an object, weight is the gravitational force acting on a body. However, for trading purposes weight is often taken to mean the same as mass.

## 1.2 Realisation of the unit of mass

The SI unit of mass is realised by linking it to the fixed numerical value of the Planck constant, this is accomplished using experiments which realise a kilogram from fundamental measurement of SI base unit and related to the fixed value of the Planck constant. units. At present there are two independent experiments which are able to realise the kilogram at the level of uncertainty required, these are the Kibble balance (Figure 1) and the X-ray crystal density (XRCD) method (Figure 2).



*Figure 1: The NPL Mk. II Kibble Balance (now at NRC in Canada)*

The Kibble Balance works by balancing a physical mass against an electro-magnetic force, thus equating mechanical power to electrical power and linking mass directly to electrical standards realised with relation to the Planck constant.

The XRCD Method, also known as the Avogadro Experiment, uses a sphere of isotopically purified silicon-28 to count the exact number of atoms in a 1 kg sphere. By knowing the mass of a single atom, the lattice spacing (how far the atoms are apart) and the volume of the sphere the total mass of the sphere can be determined.



*Figure 2: An XRCD (Avogadro) sphere on a mass comparator*

### 1.3 Kilogram consensus value

About 15 National Measurement Institutes (NMIs) around the world have Kibble Balance or XRCD experiments at various stages of maturity and levels of accuracy. A periodic comparison of the experiments is coordinated by the international standards laboratory the Bureau International des Poids et Mesures (BIPM) as part of its remit to ensure the global uniformity of measurements.

Recent results of the comparison of realisation experiments have show some discrepancy between the values for 1 kilogram produced by the various experiments so, in order to maintain the consistency of the global mass scale, a consensus value for the kilogram has been adopted. The consensus value for the kilogram is essentially an average value for all the realisation experiments which are operating at a suitable level of accuracy. The kilogram consensus value is maintained by the BIPM who continue to provide traceability for all the NMIs around the world.

## 1.4 National mass standards

About 60 countries hold platinum-iridium prototype kilograms which were initially made as copies of the IPK. The values of these standards are determined by the BIPM with relation to the consensus value of the kilogram. The National Physical Laboratory (NPL) holds the UK copy (No. 18), shown in Figure 3, which is referred to as the national prototype kilogram, or simply kilogram 18 and is the basis of the entire mass scale in the UK. The NPL participates in a wide range of international comparisons to ensure that measurements made in the UK are equivalent to those made elsewhere in the world. In the past there have been some problems with organisations based in one country not accepting traceability to any NMI other than their own. This situation has been addressed with the advent of a structured approach to international equivalence via a Mutual Recognition Agreement (MRA) and regular international measurement comparisons.



*Figure 3: The UK national standard kilogram*

The use of appropriate mass standards and their correct treatment at all times is essential to mass metrology. Weights are divided into classes from high quality reference standards (Class E<sub>1</sub>) to those used in industrial settings (Class M<sub>3</sub>). These classes were originally specified for legal metrology purposes, but they are now in common usage throughout mass measurement. The International Organisation for Legal Metrology (OIML) document OIML R111 [1] specifies the properties of weights of each class and specifies the tests that are necessary prior to carrying out a mass calibration.

## Chapter 2

# Weights

- Material
- Surface condition
- Magnetism
- Weight cleaning
- Weight handling
- Clean surfaces
- Weight storage

## 2.1 Material

Weights should be made of a material that is chemically inert, non-magnetic, hard enough to resist scratching and of a density that meets the OIML R111 recommendations for its class. Austenitic stainless steel is generally used in the construction of Class E<sub>1</sub> and E<sub>2</sub> weights. Lower accuracy weights may be manufactured from plated brass, iron or other suitable materials. Un-plated brass should be avoided due to its susceptibility to atmospheric induced surface instability.

Class E<sub>1</sub> and E<sub>2</sub> weights must be integral in construction, i.e. be made of a single piece of material, as seen in Figure 4. Other weights classes can be made up of multiple pieces with a sealed cavity to allow for adjustment. As with the other properties of weights the shape of weights for particular classes is defined in OIML Recommendation R111.



*Figure 4: A stainless steel secondary standard weight set*

## 2.2 Surface condition

Prior to use weights should be inspected for surface damage such as scratching or contamination. It is almost inevitable that weight surfaces will become slightly scratched in use, but weights with gross scratching can potentially become unstable due to contamination filling the scratch mark.

## 2.3 Magnetism

There are two magnetic properties that must be measured to characterise weights, permanent magnetisation and magnetic susceptibility. Magnetic susceptibility is a measure of whether the weight can become magnetised by being placed in a magnetic field (magnetism of this sort is transient) while permanent magnetism is a feature of a weight that cannot be altered. It may be possible to de-Gauss magnetically susceptible weights using a commercial de-Gausser, but such treatment has no effect on permanent magnetisation. OIML Recommendation R111 sets out permissible limits for these two properties for various classes of weights.

## 2.4 Weight cleaning

Stainless steel weights (OIML E and F Classes) should routinely be dusted before use using a clean soft haired brush. With the exception of this, the cleaning of weights should be avoided unless it is absolutely unavoidable as it will affect their calibration history. Where cleaning is unavoidable, it is recommended that weights should be calibrated both before and after cleaning. Rubbing with a soft, clean cloth is often enough to remove marks. Solvent cleaning should be used as a last resort. In general this will take the form of wiping the weight with a clean cloth that has been soaked in solvent.

After cleaning it is necessary to allow weights to stabilise before calibrating them. The stabilisation time will vary, from a few hours to several days, according to the class of weight and the extent of cleaning that has been undertaken [2].

Cast iron weights, if in good condition, may be cleaned by brushing with a stiff brush. Rust can be removed with a wire brush.

## 2.5 Weight handling

Weights must always be handled with the greatest care. It is important that they are never:

- touched with bare hands
- handled with sharp or abrasive tools and materials
- in contact with tools or surfaces that are not scrupulously clean
- slid across surfaces
- knocked together
- breathed on or spoken over by the operator
- cleaned by inappropriate means.

The following handling methods are recommended in order to avoid the problems mentioned above. Figure 5 illustrates good and bad practice in handling weights.

### **2.5.1 Gloves**

Gloves should be worn whenever practicable. Chamois leather is an ideal material for such gloves as it has good thermal insulation properties (so reducing thermal influences on the balance during the weighing process) and affords a good grip on large weights when manipulating them.

### **2.5.2 Tweezers**

Tweezers, forceps or other specialist lifting devices should be used whenever practicable to pick up weights and manipulate them inside the balance. It is important that these tools have no sharp edges and should, when possible, not have metal to metal contact with the weights. Metal tweezers can be covered in chamois leather to prevent metal on metal contact.

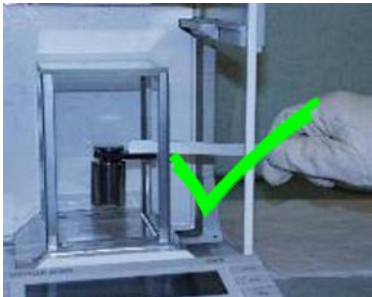


Ungloved hand, potential heat transfer to weight and enclosure.



User's hand in balance enclosure affecting the thermal stability.

Gloved hand outside enclosure.



Gloved hand outside enclosure.

**Recommended practice.**

*Figure 5: Good practice in weight handling*

## 2.6 Clean surfaces

Surfaces in a mass laboratory should be kept clean and dust free at all times. However, it is advisable to place acid free tissue paper, or something similar, on any surface prior to putting weights on it (multiple layers of tissue are recommended for cushioning uneven or hard surfaces).

## 2.7 Weight storage

Weights should be stored in a manner that protects them from damage or contamination and allows them to be easily identified. Manufacturers generally supply weights in felt lined wooden boxes. Such boxes normally have a recess for each weight and a lid which touches the weights when closed to hold them in place. Smaller weights tend to be loosely confined within a small aperture in their box.

The major disadvantage with using boxes for storage is that they contact virtually the entire surface of the weight so potentially causing contamination problems. It is important that the interiors of weight boxes are kept clean to prevent a build up of dust and other particles.

It is preferable for weights to be stored on acid free tissue paper under glass domes. This is an ideal situation as only the base of the weight is in contact with other surfaces.

Regardless of how the weight is stored it is important to ensure that it is clearly identified at all times so that it cannot get confused with similar weights.

## Chapter 3

# Weighing techniques

- Direct reading measurements
- Weighing by differences
- Substitution weighing
- ABA... calibration
- ABBA calibration
- Cyclic weighing
- Weighing by sub-division
- Calibration intervals
- Scale errors and the use of makeweights

There are many weighing techniques currently employed in mass metrology. This is indicative of the wide range of processes that rely on weighing and the uncertainty demands that are required by different industrial sectors and end-users.

### 3.1 Direct reading measurements

The most simple method of weighing is to simply place a test piece on a mass balance, such as the one shown in Figure 6, and take the displayed reading as its weight. This type of measurement is only suitable for low accuracy applications but even in this most straightforward application it is essential to follow good practice.

As with any form of mass measurement, it is essential to have the balance calibrated on a

regular basis. It is recommended that the balance undergoes a full assessment and calibration, by a suitably accredited body, on a periodic basis which will be influenced by the application and frequency of use. It is important that the balance scale is zeroed (or tared) before use. If there is an internal balance calibration feature this should be used prior to making measurements on the balance.

A reading with zero load should be taken ( $z_1$ ) followed by the reading with the load on the pan ( $r_1$ ) and a final zero reading ( $z_2$ ). This allows the user to compensate for any drift in the instrument. The drift corrected reading ( $r_d$ ) is given by:



Figure 6: Electronic balance

$$r_d = r_1 - \frac{z_1 + z_2}{2}$$

### 3.2 Weighing by differences

This technique is particularly common in analytical chemistry. In general, it involves placing a container on the balance pan, noting the reading, and then adding a substance to it. The final balance reading is noted and the difference between the two readings is taken to be the amount of material in the container. This approach is fine for relatively low accuracy requirements but is not ideal for more demanding measurements. This type of measurement is susceptible to problems with temporal balance drift (which is generally

more of a problem with modern electronic analytical balances than was the case with mechanical balances).

A more robust method of carrying out measurements of this type is to have two similar containers, one to fill with the test substance and one to act as a reference. The weighing should then be carried out as follows:

- a) The reference container is placed on the balance pan and the reading noted ( $ref_1$ )
- b) The empty container (to be loaded with the test substance) is placed on the balance ( $test_1$ )
- c) The container is filled with the test substance and the balance reading taken ( $test_2$ )
- d) The reference container is put back on the pan and the reading taken ( $ref_2$ ).

The weight of substance added to the container ( $w_t$ ) may then be calculated as follows:

$$w_t = (test_2 - test_1) - (ref_2 - ref_1)$$

This weighing scheme eliminates drift and should allow better weighing uncertainties to be achieved.

### 3.3 Substitution weighing

High accuracy weighing is generally carried out by comparing the test-weight with mass standards of similar nominal value. This is known as substitution weighing.

When weight  $A$  (the first weight to be placed on the balance pan) is compared with weight  $B$  (the second weight used) there will generally be a difference ( $\Delta m$ ) between the readings with each of the weights on the pan. This may be expressed in terms of the following equation:

$$A = B + \Delta m$$

It is important to follow a consistent nomenclature when carrying out weighings and expressing the comparison in terms of a simple equation is perhaps the best way to do this. Throughout this section the concept of an equation is used to describe a comparison weighing.

If the test-weight is of an unusual nominal value it will be necessary to calibrate it against a standard made up of several calibrated weights, or to use additional weights with the test-weight to make it up to a nominal mass. In such cases, it is recommended that a initial

approximate reading be taken as a check to ensure all the weights used have been noted down.

The number of comparisons carried out in order to calibrate a weight, and the number of independent mass standards used will depend on the uncertainty requirements of a particular calibration.

There are two popular techniques for comparative calibration: ABA... and ABBA calibrations.

### 3.4 ABA... calibration

This calibration involves the comparison of two weights, normally an unknown and a standard, by placing each one in turn on the balance pan and noting the reading. The process is symmetrical in that the weight that is placed on the pan first is also on the pan for the final reading. It is possible to vary the number of applications of each weight according to the application, but three applications of the test weight and two of the standard would be a typical regime. This leads to the following measurement scheme:

$$A_1 B_1 A_2 B_2 A_3$$

where  $A_i$  and  $B_i$  represent the balance reading for the  $i^{\text{th}}$  application of weights  $A$  and  $B$  respectively.

The difference between weights A and B is then calculated from the equation:

$$\Delta m = \frac{A_1 + A_2 + A_3}{3} - \frac{B_1 + B_2}{2}$$

This may be generalised for an ABA comparison involving  $n$  applications of weight B to become:

$$\Delta m = \frac{\sum_{i=1}^{n+1} A_i}{n+1} - \frac{\sum_{i=1}^n B_i}{n}$$

### 3.5 ABBA calibration

The ABBA calibration method is potentially more efficient than the ABA method in terms of the number of weight applications required to produce a calibration result of a particular uncertainty. In this case the weights are applied to the balance pan in the following order:

$$A_1 B_1 B_2 A_2$$

Where  $A_n$  and  $B_n$  represent the balance reading for the  $n^{\text{th}}$  application of weights A and B, respectively.

It is important in this weighing method that weighings  $B_1$  and  $B_2$  must be independent of each other. To achieve this on a typical electronic comparator the weight must be removed from the pan between readings. If a balance with an arrestment mechanism is being used it is usually sufficient to arrest the balance and open and close the balance case between the measurements. The mass difference ( $\Delta m$ ) is calculated from the following equation:

$$\Delta m = \frac{A_1 + A_2}{2} - \frac{B_1 + B_2}{2}$$

### 3.6 Cyclic weighing

High accuracy (OIML Class E and F) mass calibration usually involves calibration against two or more mass standards. Often it is necessary to calibrate several weights of the same nominal value. In this case it is inefficient to calibrate each weight against two standards. A technique known as cyclic weighing is used to rationalise the number of measurements necessary.

Comparison	Measured Difference	Corrected Difference	Using S1 value	Using S2 value	Measured value
S1 v T1	$d_1$	$d_1 - C$	$T1 = S1 - d_1 + C$	$T1 = S1^* - d_1 + C$	The average of the values for T1 to T6 calculated in the previous 2 columns
T1 v T2	$d_2$	$d_2 - C$	$T2 = T1 - d_2 + C$	$T2 = T1 - d_2 + C$	
T2 v T3	$d_3$	$d_3 - C$	$T3 = T2 - d_3 + C$	$T3 = T2 - d_3 + C$	
T3 v S2	$d_4$	$d_4 - C$	$S2^* = T3 - d_4 + C$	$S2 = T3 - d_4 + C$	
S2 v T4	$d_5$	$d_5 - C$	$T4 = S2^* - d_5 + C$	$T4 = S2 - d_5 + C$	
T4 v T5	$d_6$	$d_6 - C$	$T5 = T4 - d_6 + C$	$T5 = T4 - d_6 + C$	
T5 v T6	$d_7$	$d_7 - C$	$T6 = T5 - d_7 + C$	$T6 = T5 - d_7 + C$	
T6 v S1	$d_8$	$d_8 - C$	$S1 = T6 - d_8 + C$	$S1^* = T6 - d_8 + C$	
Sum	$d_1 + d_2 + \dots + d_8$	0			
Correction	$C = \text{sum}/8$				

Table 1: Recording the results of a cyclic weighing scheme

\*The value for S2 calculated using S1 (and vice versa) is checked against the true value of S2

Table 1 shows a typical weighing scheme for a cyclic weighing. Although two standards are used the number of weighings required is reduced from 12 to 8 through the use of this scheme. The weighing scheme is symmetrical and, as each weight is used the same number of times, the sum of the measured weight differences should be zero. In practice this is not usually the case and a correction, C, is calculated to allow for temporal changes. The weighings must be carried out in one series to allow this correction to be made. The values for the test weights T can be calculated in turn from each of the two standards and an average taken. The scheme contains two checks on the goodness of the weighings, the size of the correction C and the agreement between the values of T calculated from the two standards.

### 3.7 Weighing by sub-division

Sub-division is used for the most demanding applications. It involves the use of standards of one or more values to assign values to weights across a wide range of mass values. A typical example of this would be to use two or three 1 kg standards to calibrate a 20 kg to 1 mg weight set. Typically, this is used for OIML E Class calibrations.

This is most easily illustrated by considering how values would be assigned to a weight set using a single standard. In reality the weighing scheme would be extended to involve at least two standards of the same nominal value. The standard(s) can be compared with any

weights from the set of the same nominal value and also with various combinations of weights from the set that sum to the same nominal value. A check-weight, which is a standard treated in the same manner as any of the test-weights, is added in each decade of the calibration so that it is possible to verify the values assigned to the weight set by comparing the value measured for the check-weight with its known value. In the case of a 1 kg to 100 g weight set the following minimal weighing scheme may be used:

$$1000 = 1000S$$

$$1000 = 500 + 200 + 200D + 100C$$

$$1000S = 500 + 200 + 200D + 100C$$

$$500 = 200 + 200D + 100C$$

$$500 = 200 + 200D + 100C$$

$$200 = 200D$$

$$200 = 100 + 100C$$

$$200D = 100 + 100C$$

$$100 = 100C$$

In a simple case such as this it is possible to calculate the values of the test weights manually, but in a realistic situation when several standards are used, more comparison are made, and information is required about the weighing scheme uncertainty, it is necessary to undertake a least squares analysis using a computer.

The sub-division weighing scheme has the following advantages:

- a) It minimises use on (and hence wear on) standards
- b) It produces a set of data which provides important statistical information about the measurements and the day-to-day performance of the individual balances
- c) There is a redundancy of data (i.e. more measurements than unknowns).

It has the following disadvantages:

- a) It requires a relatively complex algorithm to analyse the data
- b) It necessitates placing groups of weights on balance pans (this can cause problems for instruments with poor eccentricity characteristics or automatic comparators designed to compare single weights).

## 3.8 Calibration intervals

The period between each calibration depends on the amount and type of use the weights experience and the accuracy class of the weights. As a general rule new weights should be calibrated annually until a reasonable calibration history (at least three calibrations) has been build up. After this, depending on the stability of the weights, the calibration period

can be extended to two or even four years. However, re-calibration may be necessary more frequently for example if the weights are dropped, contaminated or damaged.

### 3.9 Scale errors and the use of makeweights

Makeweights are weights that are added to a load in order to make it approximately equal in weight to the object it is being compared with. Makeweights are added so that only a small part of the comparator's scale is used during a comparison. For example, if a 100 g scale had a 1 % error it would equate to a 0.2 g error if there were a 20 g difference between the loads under test, but this could be reduced to a 0.01 g error if a makeweight is used to balance the loads to within 1 g.

If a makeweight of value,  $M_w$ , is added to load  $A$  during a comparison the weighing equation becomes:

$$A + M_w = B + \Delta m$$

So, in order to calculate the value of  $A$  from  $B$ , it is necessary to subtract the value of any makeweights used in association with  $A$  from the sum of the value of  $B$  and the calculated mass difference. Similarly, if makeweights had been used with weight  $B$  the value of the makeweights would be subtracted from the value of  $A$  before taking the difference in the balance readings into account.

## Chapter 4

# Air density measurement and buoyancy correction

- True mass
- Conventional mass
- Buoyancy correction
- Determination of air density from parametric measurements
- The application of buoyancy correction

The measurement of air density is necessary in the field of mass measurement to allow buoyancy corrections to be made when comparing weights of different volume in air. It is particularly important when comparing weights of different materials or when making mass measurements to the highest accuracy.

## 4.1 True mass

The mass of a body relates to the amount of material of which it consists. In terms of the calibration of weight it is referred to as true mass in order to differentiate it from conventional mass, which is generally used to specify the value of weights (see below). The international prototype of the kilogram for which the mass scale throughout the world is realised is defined as a true mass of exactly 1 kilogram. All high accuracy comparisons should be performed on a true mass basis (including class E<sub>1</sub> and E<sub>2</sub> calibrations) although values may be converted to conventional mass when quoted on a certificate.

## 4.2 Conventional mass

This is the value normally quoted on a certificate and is the conventional value of the result of a weighing in air, in accordance with International Recommendation OIML R 33. For a weight at 20 °C, the conventional mass is the mass of a reference weight of a density of 8000 kg m<sup>-3</sup>, which it balances in air of a density 1.2 kg m<sup>-3</sup>. A conventional mass value for an artefact can be calculated from a true mass value using the following equation:

$$M_c = M_t \left( 1 + \left( \frac{1}{8000} - \frac{1}{\rho} \right) 1.2 \right)$$

where:

- $M_c$  is the conventional mass of the artefact in grams
- $M_t$  is the true mass of the artefact in grams
- $\rho$  is the density of the artefact in kg m<sup>-3</sup>.

## 4.3 Buoyancy correction

This is a correction made when comparing the mass of artefacts of different volumes. It is equal to the difference in the volumes of the artefacts multiplied by the density of the medium in which they are compared (usually air). When comparing true mass values the

buoyancy correction to be applied between two artefacts can be given by the following equation:

$$BC = (V_1 - V_2) \times \rho_{air}$$

where:

- $BC$  is the buoyancy correction to be applied
- $V_n$  is the volume of the artefact,  $n$  (i.e.,  $\frac{M_n}{\rho_n}$ )
- $\rho_{air}$  is the density of the air at the time of comparison.

The calculated buoyancy correction should be applied in the form:

$$M_{t1} = M_{t2} + BC$$

Where  $M_{tn}$  is the true mass of artefact  $n$ .

When calibrating standard weights comparisons are normally performed on a conventional mass basis. For such comparisons the buoyancy correction depends on the difference in air density from the conventional value of  $1.2 \text{ kg m}^{-3}$ . Because of the way the conventional mass is specified, comparisons made in air of density exactly  $1.2 \text{ kg m}^{-3}$  require no buoyancy correction no matter what the volume difference of the weights being compared. For comparisons performed on a conventional mass basis the buoyancy correction is given by the following equation:

$$BC = (V_1 - V_2) \times (\rho_{air} - 1.2 \times 10^{-3})$$

The buoyancy correction is applied with the same convention as for the true mass correction, i.e.:

$$M_{c1} = M_{c2} + BC$$

The equation for conventional mass buoyancy correction is more complicated than for true mass and care must be taken with the sign of the correction. If weight 1 has a greater volume than weight 2 and is compared in air of density greater than  $1.2 \text{ kg m}^{-3}$  the correction ( $BC$ ) will be positive.

## 4.4 Determination of air density from parametric measurements

The standard method for determining air density involves the measurement of temperature, pressure and humidity. From these measurements and taking into account carbon dioxide concentration for the best accuracy, the density of the air can be calculated. The empirical formula for the calculation of air density recommended by the *Comité International des Poids et Mesures* (CIPM) was derived by Giacomo [3] and modified by Davis [4]. Table 2 shows typical and best achievable uncertainties for the calculation of air density from the above parameters using the CIPM formula.

	Routine measurement		Best capability	
	Uncertainty	ppm	Uncertainty	ppm
Temperature (°C)	0.1	360	0.01	36
Pressure (mbar)	0.5	500	0.05	50
Humidity (%rh/°C dew pt.)	5 %	350	0.25 °C	58
CO <sub>2</sub> content (ppm)	-	-	50	21
CIPM equation		100		100
<b>Total (x 10<sup>-3</sup> kg m<sup>-3</sup>)</b>	<b>0.86</b>	<b>720</b>	<b>0.16</b>	<b>133</b>

Table 2: Routine and best achievable realisation of air density using the CIPM formula

## 4.5 The application of buoyancy corrections

Table 3 shows the magnitude of the buoyancy correction when comparing weights of stainless steel with those of another material in air of standard density (1.2 kg m<sup>-3</sup>) on a true mass basis.

Material compared with stainless steel	Buoyancy correction (ppm)
Platinum iridium	94
Tungsten	88
Brass	8
Stainless steel	7.5*
Cast iron	24
Aluminium	294
Silicon	365
Water	875

Table 3: Buoyancy corrections when comparing dissimilar materials in air

\*This is the result of comparing two types of stainless steel, with densities 7.8 and 8.2 g cm<sup>-3</sup>

Values in Table 3 show that even when comparing weights of nominally the same material (such as stainless steel) attention must be paid to buoyancy effects when the best uncertainty is required.



*Figure 7: Stainless steel and platinum kilograms*

When comparing weights of dissimilar materials the effect of air buoyancy becomes more significant and must be applied even for routine calibrations when true mass values are being measured. Figure 7 illustrates the difference in volume for weights of the same mass manufactured from different materials.

When working on a conventional mass basis the buoyancy corrections become smaller.

International Recommendation OIML R 33 use a range for air density of  $1.1 \text{ kg m}^{-3}$  to  $1.3 \text{ kg m}^{-3}$  (ie. approximately  $\pm 10\%$  of standard air density) meaning the corrections are about one tenth of the true mass corrections. This together with the limits specified by OIML R 33 for the density of weights of Classes  $E_1$  to  $M_3$  mean that the maximum correction for any weight is one quarter of its tolerance. This is generally not significant for weights of Class  $F_1$  and below (although allowance should be made for the uncertainty contribution of the un-applied correction) but for Class  $E_1$  and  $E_2$  weights buoyancy corrections need to be applied to achieve the uncertainty values required for weights of these Classes.



## Chapter 5

# Balance assessment

- Balance location
- Management equipment
- Two-pan balances
- Single-pan mechanical balances
- Single-pan electronic balances
- Assessment of two-pan balances
- Assessment of mechanical single-pan balances
- Assessment of electronic balances
- Periodicity of calibration

The need for regular and appropriate assessment of balances is vital both in providing traceability of results and in providing accurate and reliable results. Balances can be used either as comparators or as direct reading devices. The assessment of the balance will depend on the mode in which it is used. When used as a comparator traceability for the calibration will be provided by the mass standard against which the unknown is compared and the performance of the balance is effectively given by its repeatability. When used as a direct reading device traceability for the measurement is in effect provided by the balance itself and a number of other factors such as the linearity, stability and temperature sensitivity of the balance become important. In all cases it is vital that:

- The balance assessment reflects the way in which the balance is used in practice
- It takes place in the same conditions as the balance would normally be used.

It is essential that both these conditions are met if the results of the balance assessment are to give an accurate reflection of the performance of the balance and therefore provide a valid contribution to an uncertainty budget for measurements performed on the balance.

## 5.1 Balance location

To perform at its best the balance should be placed on a level surface, located in a position as free from vibration, thermal instability, direct sunlight and draughts as possible. Care should also be taken to avoid external influences such as magnetic fields and radio frequency interference. The performance of the balance depends to a great extent on the environment in which it is located so some thought should be given to optimising its position with respect to the above parameters while allowing practical use of the balance. As previously stated, the balance should be assessed in the location in which it will be used. If a balance is used in a shop floor environment it should not be assessed in a temperature-controlled laboratory environment.

## 5.2 Measuring equipment

A basic assessment of the performance of a balance for repeatability, hysteresis and eccentric loading can be performed with two un-calibrated weights of similar nominal values. A more comprehensive assessment, however, requires a set of calibrated weights and a temperature probe to monitor the temperature of the balance. Since the sensing elements of most balances are temperature sensitive it is vital to keep a record of the balance temperature while it is being assessed as unusual variations in temperature may result in poor performance. A calibrated set of weights allow the scale error of the balance to be assessed.

## 5.3 Two-pan balances

These balances, shown in Figure 8, are now rare. They consist of a symmetrical beam and three knife edges. The two terminal knives support the pans and a central knife edge acts as a pivot about which the beam swings. Two-pan balances are generally un-damped with a “rest point” being calculated from a series of “turning points”. Some balances incorporate a damping mechanism (usually mechanical or magnetic) to allow the direct reading of a “rest point”.

Readings from this type of balance tend to be made using a simple pointer and scale although some use more complicated optical displays. In all cases the reading in terms of scale units needs to be converted into a measured mass difference.



*Figure 8: Two-pan mechanical balance*

Due to the construction of the balance, they are always used as comparators (comparing the load on one pan with the other) but the way in which the comparison is achieved can vary — double-double, double, double-substitution and substitution modes can all be used. It is vital that the balance is assessed using the same mode of comparison as is used for the normal operation of the balance.

## 5.4 Single-pan mechanical balance

These balances, shown in Figure 9, are also now rare and generally consist of a beam with two knife edges, one to support the weighing pan and one acting as a pivot. A fixed counterweight balances the load on the pan.

This type of balance is usually damped and a series of built in weights allow the calibration of a range of mass values by maintaining a constant load on the balance beam.

The displays on these balances tend to be of the optical variety and the sensitivity of the balance can usually be adjusted by the user.



*Figure 9: Single-pan mechanical balance*

## 5.5 Single-pan electronic balances

These are usually top loading balances with the applied load being measured by an electromagnetic force compensation unit or a strain gauged load cell. Single pan electronic balances give a direct reading of the weight applied whereas the other two mechanical balance types rely on the comparison of two forces (an unknown weight with either an external or internal weight). Despite the possibility of using these balances as direct reading devices (applying an unknown weight and taking the balance reading as a measure of its mass) single pan electronic balances will always perform better when used as comparators, comparing a standard and an unknown in an ABA or ABBA sequence. As with all the balance types discussed the method of assessment should reflect the way in which the balance is used in practice.

## 5.6 Assessment of two-pan balances

This type of balance is less frequently used, mainly due to the amount of time it takes to make a weighing compared with electronic balances. For this reason, only a brief outline of the tests that should be performed on this type of balance is given. Further details of these tests may be found in NPL Notes on Applied Science No.7, 'Balances, Weights and Precise Laboratory Weighing', 1954.

### 5.6.1 Routine periodic checks

The following tests should be carried out on a regular basis and are essential to the routine operation of the balance.

### 5.6.2 Rider weight

The effective value (mass value corrected for air buoyancy based on the actual density of the weight) of the rider weight should be measured against suitable mass standards.

### 5.6.3 Internal balance weights

Instead of (occasionally as well as) a rider bar some two-pan balances have built in poising weights. These should be removed and calibrated in the same way as the rider weight.

### 5.6.4 Linearity of scale

This can be checked using either fractional weights or with the rider and rider bar.

### 5.6.5 Sensitivity

Measuring the sensitivity of the balance allows a measured difference in terms of scale divisions to be converted into a mass difference and is therefore vital to the use of the balance. A suitable sensitivity weight, which can be easily transferred between the balance pans without arresting the balance, is required. The value of this weight should be enough to cause the pointer to move along the scale by between one quarter and one third of the total scale length. The sensitivity weight should be calibrated against suitable mass standards. The weight is then swapped between the balance pans (left to right and then right to left) without arresting the balance. Arresting the balance would change the effective rest point slightly and compromise the accuracy of the sensitivity measurement. The average effect of swapping the sensitivity weight ( $d_{av}$ ) is calculated (the two figures should agree to better than 5 %) and this value used to calculate the sensitivity as follows:

$$S = \frac{2 \times W_s}{d_{av}}$$

where:

- $d_{av}$  is the average scale difference
- $S$  is the measured sensitivity
- $W_S$  is the effective value of the sensitivity weight

The sensitivity should be checked at at least four points across the weighing range of the balance.

### 5.6.6 Repeatability of reading

The repeatability of reading can be checked by loading the balance and performing a series of releases calculating a rest point for each series. Apart from arrestment the balance remains undisturbed between releases making this measurement relatively insensitive to the user. This makes this test useful in providing a temporal measurement of the balance performance which is insensitive to external influences.

### 5.6.7 Repeatability of measurement

This provides a more accurate assessment of the balance's performance "in use". A series of repeated weighings are performed using the balance in its normal comparison mode (double weighing or substitution). The series of measured mass differences can be statistically analysed to give a measure of the balance performance.

The average measured mass difference can also be compared with the certified mass difference between the weights taking into account the uncertainties on these certified values.

## 5.7 Single-pan balances

Single-pan mechanical balances have generally been replaced by electronic balances, which often offer better resolution and are easier to use. The range of tests that should be performed on these balances is only briefly described.

### 5.7.1 Visual inspection and mechanical check

- Check knife-edges and planes
- Ensure the released balance does not foul
- Adjust the zero of the balance.

### 5.7.2 Drift

This will give an indication of the balance's stability and its sensitivity to changes in environmental conditions. It will also give an indication of how long the balance should be left to give a stable reading (the effectiveness of the damping). If the drift is linear it may be eliminated by the use of a suitable symmetrical weighing method (e.g. ABA weighing).

### **5.7.3 Calibration of internal weights**

Ideally the weights built into the balance should be removed and calibrated externally. If this is not possible, they can be left in the balance and calibrated by dialling them up in combinations.

### **5.7.4 Effect of off-centre loading (eccentricity)**

The effect of off-centre loading on the balance reading is assessed using a weight of nominal value equal to about half the balance capacity (to avoid mechanical damage to the balance). The weight is placed at the extremes of the pan and the results compared with the reading when the weight is placed centrally. A significant difference between the reading at the centre and extremes of the pan may indicate the balance is fouling. This test is identical to that described in more detail for a single-pan electronic balance.

### **5.7.5 Hysteresis**

This is checked by taking readings of increasing and decreasing load. Significant difference may indicate mechanical problems within the balance. This test is identical to that described in more detail for a single pan electronic balance.

### **5.7.6 Scale sensitivity (scale value)**

This is checked by placing a weight on the pan equal to the full range value of the (optical) scale. For most single-pan mechanical balances adjustment of the scale sensitivity is possible. This test is identical to that described in more detail for a single-pan electronic balance.

### **5.7.7 Scale linearity**

The (optical) scale should be checked for linearity at regular intervals along its range using a calibrated set of (fractional) weights. This test is identical to that described in more detail for a single-pan electronic balance.

### **5.7.8 Repeatability of reading**

The repeatability of reading should be assessed by loading the balance (usually to maximum load) and taking a number of readings. The balance should be arrested and released between readings without otherwise disturbing it. This provides a measure of the balance performance, independent of external variables, and can thus be used as a repeatable check of the balance performance over time.

### 5.7.9 Repeatability of measurement

This represents a series of actual comparisons using the balance in its normal weighing mode. Statistical analysis of the repeatability of measurement provides a practical assessment of the balance performance under normal weighing conditions.

## 5.8 Assessment of electronic balances

This type of balance, an example of which is shown in Figure 10, is by far the most widely used and the simple principle of operation means that a full assessment takes the form of relatively few tests.

As with the other types of balance the manner in which the balance is assessed will depend on the way it is to be used in practice. Before starting the assessment of an electronic balance, it should have been left switched on for a least an hour and preferably overnight (check the manufacturer's literature for details) and should be checked that it is level. Where appropriate the balance calibration facility should be used before starting the assessment.



Figure 10: Single-pan electronic balance

### 5.8.1 Hysteresis

Hysteresis occurs when, for a given weight, the balance displays a different reading depending on whether the load is increasing or decreasing. It can be checked by:

- Incrementally increasing the load by adding weights one at a time, noting the balance reading after each addition
- Remove the weights in reverse order, noting the reading after each removal
- Compare the readings for the same load being applied to the balance. Significant deviation between the increasing and decreasing readings indicates a poorly adjusted or dirty balance.

Significant deviation between the increasing and decreasing readings indicates a poorly adjusted or dirty balance.

### 5.8.2 Effect of off-centre loading (eccentricity)

Electronic balances are generally more sensitive to off centre loading than either of the other types, as there is often little de-coupling between the balance pan and the sensing element. This effect is assessed by placing a weight, of value about half the capacity of the balance, at the extremes of the pan. Compare the results with the reading when the weight is placed centrally. Typical results are shown in Table 4.

Reading sequence	Centre	Right	Front	Left	Back
C > R > F > L > B	0.86 g	0.81 g	0.83 g	0.87 g	0.82 g
B > L > F > R > C	0.85 g	0.87 g	0.80 g	0.86 g	0.83 g
Mean reading	0.855 g	0.84 g	0.815 g	0.865 g	0.825 g
Reading relative to centre		-0.015 g	-0.04 g	+0.01 g	-0.03 g

Table 4: Weighing results for an eccentricity test on a top-pan balance

The weight is placed at each position on the pan twice and the weighing scheme is symmetrical with time, so any balance drift is calculated out.

### 5.8.3 Scale error and linearity

Scale error and linearity can be checked with a suitable set of calibrated mass standards. Measurements should be made at about ten equal steps across the range of the balance. If the balance is usually used at a particular load or an area of the scale is thought to be non-linear, additional linearity checks can be performed at that load.

### 5.8.4 Repeatability of reading

The repeatability of an electronic top pan balance can be assessed by repeated application of a mass standard. A set of ten weight applications should be measured at at least two

loads (usually half and full load) and at any other load where the balance will be normally used.

### 5.8.5 Repeatability of measurement

A more relevant test for a balance that is always used as a comparator is to carry out repeated comparisons of a pair of weights using the same weighing scheme that would normally be used on the balance.

Generally, the following data should be calculated for a repeatability assessment:

- Maximum difference between consecutive weighings
- Maximum difference between any weighings
- Standard deviation of the series ( $\sigma_{n-1}$ )
- Standard deviation of the series accounting for drift.

## 5.9 Periodicity of calibration

The period between balance calibrations depends on the intensity of use of the balance and on the type of measurements which are being performed. In general, a full balance assessment should be undertaken annually and if any changes to the balance set up or location have occurred. If a balance is fitted with an internal calibration facility, this should be checked on a weekly basis or if changes are made to the balance set up.

## Chapter 6

# Uncertainty in mass calibration

- Introduction
- The measurement
- The uncertainty budget
- Air buoyancy uncertainty budget
- Reporting the results

## 6.1 Introduction

When making measurements there is always an element of uncertainty in the result. We cannot know ‘true’ values — there are limitations in our knowledge and in the performance of the instruments we are using. Therefore, a measurement is not complete without an estimate of the doubt that surrounds it (the uncertainty) and the confidence we have in that estimate. This chapter gives an introduction to the terminology and the main sources of uncertainty in the calibration of a weight of nominally 100 g, when using a comparator, plus a brief summary of the process for calculating the uncertainty and associated confidence level. All calculations are performed in accordance with the ISO Guide to the Expression of Uncertainty in Measurement [5] also known as the GUM.

## 6.2 The measurement

Although perhaps an obvious point, before starting it is worth confirming precisely what the measurements are aimed at determining. In this example it is the conventional mass of a weight and the following need to be considered:

- Which measurements and calculations will be required to enable you to establish the mass value and the uncertainty in its determination? For example, will you need to determine air density?
- How many measurements do you need to take? (The more measurements you take the more representative the mean (average) value becomes, although there is a reduction in benefit as the number of measurements increase beyond a certain point. Ten measurements is a common choice and statistically valid, but not always practical — e.g., for economic reasons. One way of dealing with this problem is to use some data from previous measurements to determine the performance of the balance and then take a smaller number of measurements for the particular calibration in hand. This is dealt with in paragraph 6.3.3 under *repeatability*.)
- How to take your measurements and calculate a mass value.

## 6.3 The uncertainty budget

Once you have determined the mass value you are ready to start calculating the uncertainty in the measurement. Table 6 shows a template for a typical uncertainty budget. It can be in the form of a computer spreadsheet — to make repeated calculations easier — or it can be a paper table completed by hand using a calculator. Each column in the table is dealt with separately below.

Table 5: Uncertainty budget template

Symbol	Source of uncertainty	Value	Probability distribution	Divisor	Sensitivity coefficient	Standard uncertainty	$v_i$ or $v_{eff}$

### 6.3.1 Symbol

The symbol used in to denote the input quantity or the influence factor.

### 6.3.2 Sources of uncertainty

The sources of the uncertainty are dependent on the measurement process and equation used — as defined in your procedures and by your laboratory environment. Here we consider the most common sources.

Each of the *input quantities* in the measurement equation used to calculate the mass value has an uncertainty. For example if the equation you are using is:

$$W_X = W_S + \Delta_W + Ab$$

where:

- $W_X$  is the unknown mass
- $W_S$  is the mass of the standard
- $\Delta_W$  is the difference in the balance readings
- $Ab$  is the correction for air buoyancy.

There is an uncertainty associated with each of the input quantities  $W_S$ ,  $\Delta_W$  and  $Ab$ . The uncertainty in  $\Delta_W$  depends on uncertainty due to other influence factors:

- Rounding errors in the comparator readings ( $\delta I_d$ )
- The uncertainty due to less-than-perfect repeatability of the readings ( $W_R$ )
- The value associated with comparator linearity ( $\delta C$ ).

Another influence factor to be considered is

- The drift of the standard with time ( $D_S$ ).

The uncertainty sources are used to populate the template, as shown in Figure 7.

Symbol	Source of uncertainty	Value	Probability distribution	Divisor	Sensitivity coefficient	Standard uncertainty	$v_i$ or $v_{eff}$
$W_S$	Calibration of mass standard						
$\delta C$	Comparator linearity						
$Ab$	Air buoyancy						
$D_S$	Uncorrected drift of the standard						
$\delta I_d$	Digital rounding error						
$W_R$	Repeatability						

Table 6: Uncertainty budget including uncertainty sources for a mass calibration (by comparison)

### 6.3.3 Values

The values associated with the sources of uncertainty are either measured, calculated or come from a *a priori* (previous) knowledge. Potential sources of uncertainty are illustrated in Table 8.

In our example:

Symbol	Uncertainty source	The value
$W_S$	Calibration of mass standard	The uncertainty in the mass standard is taken from its calibration certificate.
$\delta C$	Comparator linearity	Estimated from previous measurements according to your procedures.
$Ab$	Air buoyancy	A calculated uncertainty based on the air buoyancy correction equation i.e., $(V_s - V_x)(\rho_a - 1.2)$ (see paragraph 6.4) In this example the volumes have not been measured and the uncertainty in the volume difference is based on the values given in the OIML recommendation R111 (if the volumes are measured the uncertainty can be less) but the value of air density has been measured.
$D_S$	Uncorrected drift of the standard	The uncertainty quoted on the mass standard's calibration certificate will not include any contribution for drift in its mass value. The evaluation of this effect is normally the responsibility of the weight's owner as they are best placed to evaluate how much its mass changes between calibrations. Drift is usually determined by considering how much a particular artefact has changed its mass value over a recent period and extrapolating the figure to cover the period up to its next calibration. In this example no previous calibration knowledge is assumed and the uncertainty in the current mass value calibration is also used to estimate the limits of drift.
$\delta I_d$	Digital rounding error	Each reading is subject to a rounding error. It is taken to be $\pm$ half the resolution of the comparator. Such errors occur in the comparator reading of the standard mass and the unknown mass.
$W_R$	Repeatability	This is an uncertainty component which is a measure of the 'spread' of the repeated readings. It is estimated by determining the experimental standard deviation of the mean (see <i>Further reading</i> ). In this example a previous evaluation of repeatability of the measurement process (from ten comparisons between a mass standard and an unknown mass) were used to establish a standard deviation of 0.00017 mg which was then divided by $\sqrt{n}$ , where n is the number of readings in the current measurement — in this case three).

Table 7: Uncertainty sources and the values which need to be used

### 6.3.4 Probability distributions

A probability distribution is a statistical description of how results behave. There are three distributions commonly used in mass uncertainty budgets: *normal* (sometimes called *Gaussian*), *rectangular* (sometimes called *uniform*) and *triangular*. The following graphs illustrate these distributions where the zero value on the horizontal axis represents the

mean value of a number of readings, the vertical axis is the probability of a particular value occurring and the broken lines represent plus and minus one standard deviation ( $k = 1$ ) — encompassing approximately 68 % of the measurement values (i.e. 68 % of the area under each curve).

#### 6.3.4.1 Normal distribution

Illustrated in Figure 11, this represents a group of measurements where the values are more likely to fall closer to the mean value than further away from it. Repeated measurements are an example of this type of distribution. The graph shows normally distributed data with a mean value of zero and standard deviation of  $\pm 1$ .

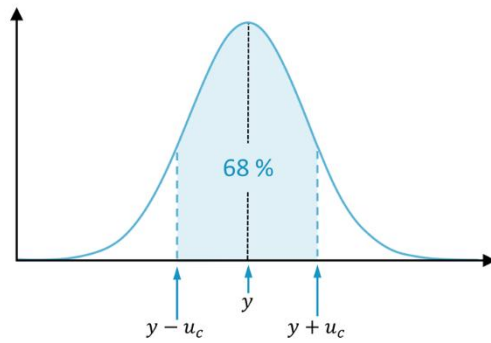


Figure 11: Typical normal uncertainty distribution

#### 6.3.4.2 Rectangular distribution

Illustrated in Figure 12, this represents a group of measurements where the values are evenly spread between two limits and never fall outside these limits. An example is when using an assumed air buoyancy correction (as opposed to a measured or calculated value). The graph shows a rectangular distribution, again with a mean value of zero, but limits of  $\pm 5$ . In this case one standard deviation is  $\pm 2.89$ .

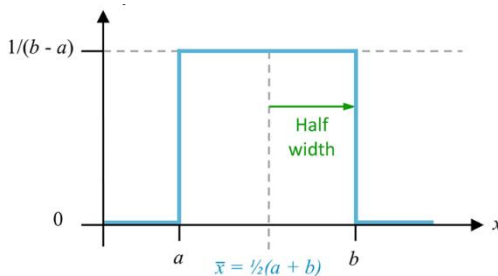


Figure 12: Typical rectangular uncertainty distribution

### 6.3.4.3 Triangular distribution

This is the distribution you get when adding two rectangular distributions. An example is the uncertainty due to rounding errors. The graph shows a triangular distribution, again with a mean value of zero and limits  $\pm 5$ . In this case, one standard deviation is  $\pm 2.04$ .

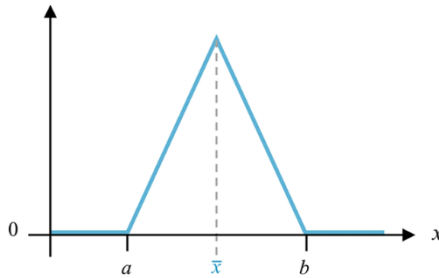


Figure 12: Typical triangular uncertainty distribution

### 6.3.5 Divisor

In order to eventually sum all the individual input quantities, they must be quoted with the same confidence level. This is done by establishing a number by which the input quantity uncertainty value is divided to convert it to one standard deviation and is dependent on the distribution as shown in Table 9.

Distribution	Divisor
Normal	1 or 2
Rectangular	$\sqrt{3}$
Triangular	$\sqrt{6}$

Table 8: Divisors for common uncertainty distributions

The normal distribution has a divisor of either 1 or 2 depending on the confidence level of the value quoted. For example, on a certificate of calibration the uncertainty might be quoted as ' $k = 1$ ' (~68 %) or ' $k = 2$ ' (~95 %) in which case the divisor is the ' $k$ ' number. Other values are sometimes used, for example  $k = 3$ .

### 6.3.6 Sensitivity coefficient ( $c_i$ )

This is a multiplication factor that converts the uncertainty in the value of an input quantity to a corresponding uncertainty in the output quantity (it sometimes has to convert both *quantity* — such as temperature or pressure — to mass, and also the correct *units*). In the example being discussed, all the input quantities are already expressed in the quantity *mass* and using the sub-unit *milligram* so the sensitivity coefficient is 1.

An explanation of the air buoyancy correction calculation is in paragraph 6.4.

### 6.3.7 Standard uncertainty in units of measurand, $u_i(W_i)$

In order to add all the components together, we need them in the same units and the values for this column are simply calculated from

$$\text{Value} \div \text{Divisor} \times \text{Sensitivity coefficient } (c_i)$$

### 6.3.8 Degrees of freedom, $\nu_i$

The number of degrees of freedom is “...in general the number of terms in a sum minus the number of constraints on the terms of the sum” [5]. Before considering this further, it is necessary to first appreciate that measurement uncertainties are considered to fall into one of two categories, known as *Type A* and *Type B*.

- Type A uncertainties are those that are evaluated by statistical methods. For example, uncertainty due to less-than-perfect repeatability of a measurement can be reduced by calculating a mean value from several measurements
- Type B uncertainties are evaluated by other means. They cannot be reduced by taking more measurements — the uncertainty quoted on a certificate of calibration cannot be made lower by repeatedly reading the certificate for example!

The degrees of freedom,  $\nu_i$ , for individual uncertainty contributions are given by:

Type A	$\nu_i = n - 1$ where n is the number of measurements used to evaluate the type A contribution
Type B	$\nu_i$ is usually taken to be infinite

Table 9: Degrees of freedom for Type A and Type B uncertainties

### 6.3.9 Adding it all up

#### 6.3.9.1 Combined standard uncertainty $u(W_x)$

To obtain an uncertainty in  $W_x$ , the mass value of the unknown weight, the components have to be added to obtain a *combined standard uncertainty*. As it is unlikely that all the errors will have been at their maximum value in any one measurement it is inappropriate to add them in arithmetically. The recognised way to address this issue is to arithmetically add the squares of the standard uncertainties and then take the square root of the result — this process is known variously as taking the *root sum of the squares* (RSS) or *quadrature summation*.

In the example we are considering the summation is:

$$\begin{aligned}
 u(W_x) &= \sqrt{0.0250^2 + 0.0115^2 + 0.0216^2 + 0.0289^2 + 0.0020^2 + 0.0001^2} \\
 &= 0.0454 \text{ mg}
 \end{aligned}$$

The resulting probability distribution will be a normal distribution unless one rectangular distribution is much larger than the other components.

### 6.3.9.2 Effective degrees of freedom $v_{eff}$

In general, the effective degrees of freedom,  $v_{eff}$ , will not need to be calculated if the type A uncertainty is less than half of the combined standard uncertainty, there is only one type A component and at least three measurements have been taken. Otherwise, the effective degrees of freedom will have to be calculated to ensure that the  $k$ -factor of 2 will indeed give a confidence level of ~95 %.

The effective degrees of freedom for the combined standard uncertainty will depend on the magnitude of the degrees of freedom for the type A contributions in relation to the type B. If the type B uncertainties are all taken to have infinite degrees of freedom the relationship is shown using the simplified Welch-Satterthwaite equation:

$$v_{eff} = \frac{u_c^4(y)}{\left(\frac{u_i^4(y)}{v_i}\right)}$$

where

$u_c(y)$  is the combined standard uncertainty

$u_i(y)$  is the individual type A uncertainty contribution

$v_i$  is the degree of freedom in  $u_i(y)$

therefore in our example:

$$v_{eff} = \frac{0.0454^4}{\left(\frac{0.0001^4}{9}\right)}$$

thus  $v_{eff} = 3.8E + 11$

In this example  $v_{eff}$  is a very large number which can be taken to be infinity. If this value had been less than 100 a  $k$ -factor would have been calculated from a distribution other than a normal distribution. More information about this can be found in the further reading list but for our example the  $k$ -factor is two.

### 6.3.9.3 Expanded uncertainty

The expanded uncertainty  $U(W_x)$  is the combined standard uncertainty  $u(W_x)$  multiplied by a  $k$ -factor which will give an uncertainty value with a confidence level of approximately 95%, in this case 2.

The final uncertainty budget for our example is shown below in Table 10.

Symbol	Uncertainty source	Value ±mg	Probability distribution	Divisor	Sensitivity coefficient	Std. uncertainty ±mg	$\nu_i$ or $\nu_{eff}$
$W_S$	Calibration of standard weight	0.0500	Normal	2	1	0.0250	$\infty$
$\delta C$	Comparator linearity	0.0200	Rectangular	$\sqrt{3}$	1	0.0115	$\infty$
$Ab$	Air buoyancy	0.0216	Normal	1	1	0.0216	$\infty$
$D_S$	Uncorrected drift of the standard	0.0500	Rectangular	$\sqrt{3}$	1	0.0289	$\infty$
$\delta I_d$	Digital rounding error	0.0050	Triangular	$\sqrt{6}$	1	0.0020	$\infty$
$W_R$	Repeatability	0.0001	Normal	1	1	0.0001	9
$u(W_x)$	Combined standard uncertainty		Normal			0.0454	> 500
$U(W_x)$	Expanded uncertainty		Normal $k = 2$			0.0908	> 500

Table 10: Fully populated mass uncertainty budget

## 6.4 Air buoyancy uncertainty budget

In order to calculate the uncertainty in the air buoyancy correction for entry into the main uncertainty budget an additional uncertainty budget has to be completed. Air buoyancy ( $Ab$ ) is dependent on the volumes of the weights and the air density with the following relationship:

$$Ab = (V_s - V_x)(\rho_a - 1.2)$$

where:

- $(V_s - V_x)$  is the difference in volume between the standard and the unknown weight
- $(\rho_a - 1.2)$  is the difference between measured density of the air and the standard air density.

There are two ways to calculate an uncertainty value for  $Ab$ , either working in relative values or calculating the sensitivity coefficient directly. In our example the volumes have not been measured so the value of  $(V_s - V_x)$  is taken to be the largest difference possible according to the OIML recommendations [1] when comparing  $E_2$  and  $F_1$  weights — that is  $1.3 \text{ cm}^3$  with an uncertainty of  $\pm 1.3 \text{ cm}^3$ . This uncertainty is treated as a rectangular distribution because the real value may lie anywhere between these limits and thus the standard uncertainty ( $u(V)$ ) is equal to  $\pm(1.3 \div \sqrt{3})$ . In this example the air density,  $\rho_a$ , has been measured as being  $1.22 \text{ kg m}^{-3}$  with an uncertainty of  $\pm 10 \%$  — thus the uncertainty in  $(\rho_a - 1.2)$ , that is  $u(\rho)$ , is  $\pm 0.012 \text{ kg m}^{-3}$ . Thus:

$$Ab = (1.3)(1.22 - 1.2) = 0.026$$

#### 6.4.1.1 Relative uncertainty

The relative uncertainty in the air buoyancy,  $\frac{u(Ab)}{Ab}$ , is given by:

$$\frac{u(Ab)}{Ab} = \sqrt{\left(\frac{u(V)}{V_x - V_s}\right)^2 + \left(\frac{u(\rho)}{\rho_a - 1.2}\right)^2}$$

Inserting the values gives:

$$u(Ab) = Ab \sqrt{\left(\frac{1.3 \div \sqrt{3}}{1.3}\right)^2 + \left(\frac{0.012}{1.22 - 1.2}\right)^2}$$

$$u(Ab) = 0.0216$$

This method of calculation works well for an equation where the only operators are multiplication or division.

#### 6.4.1.2 Partial differentiation

The other method, partial differentiation, sounds more complicated but is actually quite straightforward for this type of equation.

Using the same equation,  $Ab = (V_s - V_x)(\rho_a - 1.2)$ , and the same values as above, we calculate the sensitivity coefficients — the numbers by which values of  $u(V)$ , expressed here in  $\text{cm}^3$ , and  $u(\rho)$ , expressed here in  $\text{kg m}^{-3}$ , should be multiplied to calculate their effect on the output quantity expressed in grams.

The partial derivative of a simple equation, such as the one we are looking at, is simply the multiplier for the term for which we wish to calculate the partial derivative. For example, if our equation is  $A = B \times C$  then the partial derivative of  $B$  is  $C$  and the partial derivative  $C$  is  $B$ .

In our air density problem, the partial derivative of  $(V_s - V_x)$  is  $(\rho_a - 1.2)$  and the partial derivative of  $(\rho_a - 1.2)$  is  $(V_s - V_x)$ . In the correct terminology this is expressed as:

$$\frac{\partial(Ab)}{\partial(V_s - V_x)} = (\rho_a - 1.2) = 0.02$$

$$\frac{\partial(Ab)}{\partial(\rho_a - 1.2)} = (V_s - V_x) = 1.3$$

The values 0.02 and 1.3 are entered directly into the sensitivity coefficient column of the uncertainty budget, shown in Table 11. The remaining calculations are the same as for the main uncertainty budget.

Symbol	Source of uncertainty	Value ±mg	Probability distribution	Divisor	Sensitivity coefficient	Std. uncertainty ±mg	$v_i$ or $v_{eff}$
$V_s - V_x$	Difference in volumes	1.3	Rectangular	$\sqrt{3}$	0.02	0.0150	$\infty$
$\rho_a - 1.2$	Diff. in air density	0.012	Normal	1	1.3	0.0156	$\infty$
Combined standard uncertainty			Normal			0.0216	$\infty$

Table 11: Uncertainty contributions for air buoyancy

The combined standard uncertainty is the same as in the previous method.

Partial differentiation is more difficult when the equation is more complex.

## 6.5 Reporting the results

In general calculations should not be rounded until the final result is calculated. The uncertainty should be quoted to 2 significant figures and the result quoted to the same number of decimal places. In the worked example the result would be reported in the form:

$$100.000\ 71\ \text{g} \pm 0.10\ \text{mg}$$

and would be accompanied by a statement explaining how the uncertainty and its confidence level are calculated such as:

*The reported expanded uncertainty is based on a standard uncertainty multiplied by a coverage factor  $k = 2$ , providing a level of confidence of approximately 95%. The uncertainty evaluation has been carried out in accordance with UKAS requirements.*

This statement was taken from the UKAS document [6].

The uncertainty has been rounded to 0.10 mg; uncertainties should always be rounded up rather than down to ensure that the value remains within the 95 % confidence limit.

## Chapter 7

# References

- References

- [1] International Organisation of Legal Metrology (OIML), International Recommendation No 111:2004 Weights of classes E1, E2, F1, M1, M2, M3.
- [2] Clarkson M.T., May B.J Metrologia, **38(2)**, 161-171, 2001
- [3] Giacomo P., Metrologia, **18**, 33-40, 1982
- [4] Davis R.S., Metrologia, **29**, 67-70, 1992
- [5] *JCGM GUM-1:2023 Guide to the expression of uncertainty in measurement*
- [6] UKAS publication M3003. *The Expression of Uncertainty and Confidence in Measurement*, 2024.

